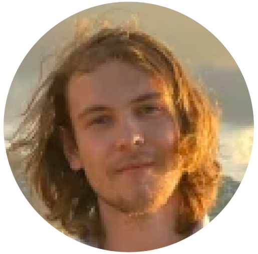
# Locally Sampleable Distributions

## Hamming slices, symmetry, and quantum advantage


**Kewen Wu** Caltech

**Daniel Grier**
UCSD

**Daniel M. Kane**
UCSD



**Jackson Morris**
UCSD

**Anthony Ostuni**
UCSD

# Plan

## Introduction and applications

Data structure lower bounds

Input-independent quantum advantage

## Proof outline

Granularity, or Hamming weight concentration

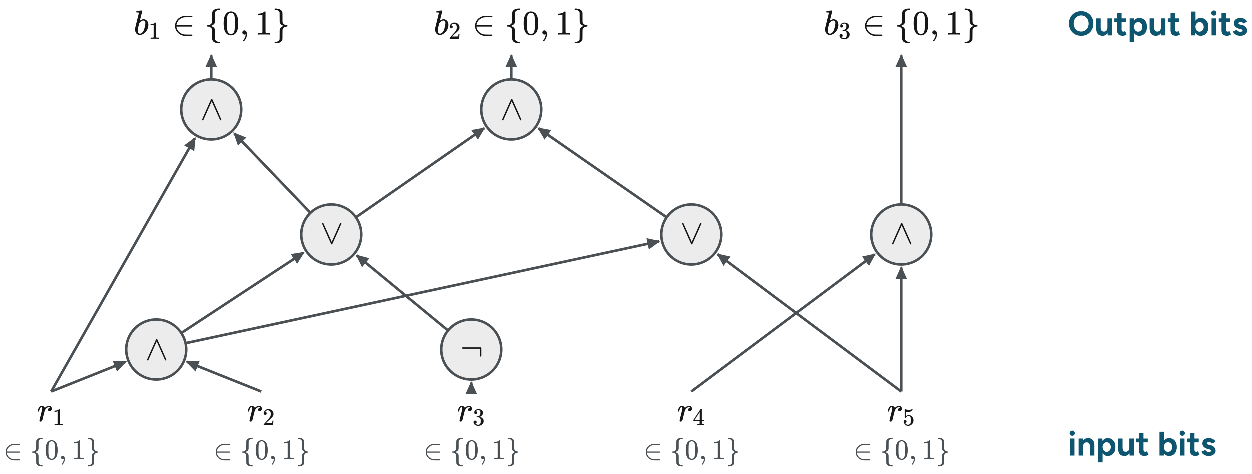
The graph pruning theorem

Classifying locally sampleable symmetric distributions

## Summary and open problems

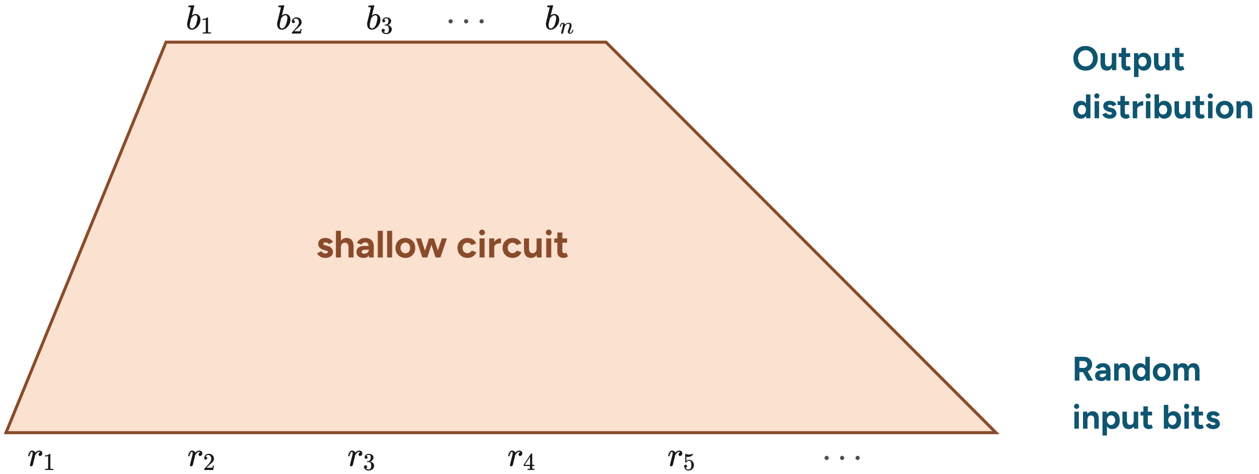
# What distributions can a shallow circuit produce?

constant depth, bounded fan-in



# What distributions can a shallow circuit produce?

constant depth, bounded fan-in




low-degree polynomials · small decision trees · weakly correlated computations

# Formal set-up

Let $f : \{0,1\}^m \to \{0,1\}^n$ be $d$-**local**

each output bit depends on at most $d$ input bits

$U$ is uniform on $\{0,1\}^m$, and $f(U)$ is the induced output distribution

$m$ is unrestricted · $f$ is deterministic · $d$ is a large constant

**The asymmetry**

Output fan-in is at most $d$, but input fan-out is unbounded

A single input bit may influence every output bit

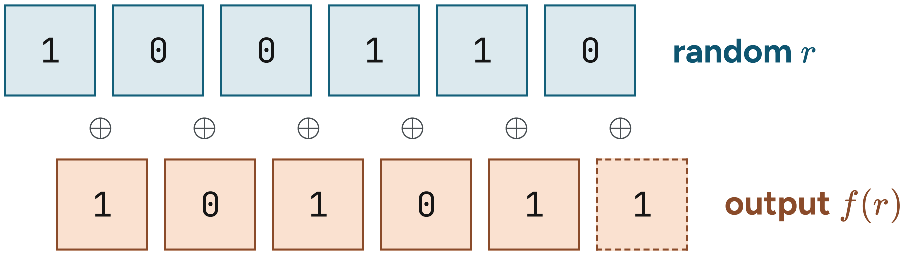
# The evens distribution

Uniform over $\{x \in \{0,1\}^n : |x| \equiv 0 \bmod 2\}$

**The natural sampler.** Draw $b_1, \ldots, b_{n-1}$ uniformly, then set $b_n = b_1 \oplus \cdots \oplus b_{n-1}$.

**Parity is not local.**
An output bit that sees only $d$ inputs cannot be the parity of $n-1$ bits.

## Can we sample the distribution? **Yes!**




$f_i(r) = r_i \oplus r_{i+1}$ cyclically

# Other ones?

Uniform over weight $\leq 10$?

Uniform over weight exactly $n/2$?

Uniform over weight $0$ modulo $3$?

Uniform over strings with more ones than zeros?

**What uniform symmetric distributions can be locally sampled?**

**Symmetric**: support is a symmetric set

**Uniform**: uniform over its support

# A classification: uniform symmetric case

| $f(U) \equiv 0^n$ | $f(U) \equiv 1^n$ | $f(U) \sim \{0^n, 1^n\}$ |
|---|---|---|
| `zeros` | `ones` | `zeros-or-ones` |
| 0-local | 0-local | 1-local |

| $\lvert f(U) \rvert \equiv 0 \bmod 2$ | $\lvert f(U) \rvert \equiv 1 \bmod 2$ | $f(U) \sim \{0,1\}^n$ |
|---|---|---|
| `evens` | `odds` | `evens-or-odds` |
| 2-local | 2-local | 1-local |

**Conjecture** [Filmus–Leigh–Riazanov–Sokolov'23]. No other examples.

**Theorem** [KOW'25]. No other examples — and this holds even approximately.

If $f(U)$ is $\varepsilon$-close to some uniform symmetric distribution, it is $O_d(\varepsilon)$-close to one of the six.

# A further classification: general symmetric case

**Theorem** [KOW'26]

If a $d$-local $f(U)$ is close to **any** symmetric distribution, it is close to a mixture of

**evens**, **odds**, and $\mathrm{Ber}(\gamma)^n$ with $\gamma$ an integer multiple of $2^{-d}$

and the mixing weights are biases of $\mathbf{F}_2$-polynomials of $O_d(1)$ degree with $O_d(n)$ monomials.

compute each monomial then take their parity,
using the same trick as for **evens**

**Consequence.** The mixture is itself an $O_d(1)$-local distribution.

# Motivations

A natural question on its own

Data structure lower bounds [Vio'12, FLRS'23, KOW'24]

Quantum–classical separation [BGK'18, WP'23, GKMOW'26]

# Data structure lower bound

**The dictionary problem**

Given an $n$-bit string $x$ of Hamming weight $0$ modulo $127$,
store it as an $s$-bit string $A(x)$ so that every $x_i$ can be recovered from *few* bits of $A(x)$.

$H = \{\, x \in \{0,1\}^n : |x| \equiv 0 \bmod 127 \,\}$ and $A : H \to \{0,1\}^s$

For each $i$, some $B_i$ reads at most $q$ bits of $A(x)$ and outputs $x_i$

*What is the best trade-off between storage $s$ and query cost $q$?*

# Two extremes

### Maximum query efficiency

Store $A(x) = x$

$s = n$

1 bit of $A(x)$ decodes $x_i$

### Maximum compression

Only $\approx 2^n/127$ possible $x$; store the index

$s = \log(2^n/127) = n - 6$

all of $A(x)$ to decode $x_i$

*Can we have both?* **No.** [KOW'24]

Either you query $\omega(1)$, or you store full length $n$.

# A local decoder is a local sampler [Viola'12]

Feed the decoders a **uniformly random** string instead of a real $A(x)$:

$$a \sim \{0,1\}^s \longrightarrow B(a) = (B_1(a), \ldots, B_n(a)) \text{ is a } q\text{-local distribution}$$

Whenever $a$ happens to equal $A(x)$ for some $x \in H$, the output is exactly that $x$:

$$\Delta_{\mathrm{TV}}\big(B(U_s),\, U_H\big) \;\leq\; 1 - |H|/2^s$$

Sampling distance lower bound → Data structure lower bound

So it is enough to show no $q$-local distribution is close to uniform on strings of weight $0 \bmod 127$.

# Quantum-classical separation

$|0\rangle^{\otimes n}$ → shallow quantum → $D_q$

random bits → shallow classical → $D_c$

*Can shallow quantum circuits produce a distribution far from all classical shallow circuits?*
*Can we prove quantum advantage in tasks where no input is required?*

**Theorem** [GKMOW'26]. There is an explicit depth-**7**, geometrically local quantum circuit with

$$\Delta_{\mathrm{TV}}(D_q, D_c) \ \geq\ 1 - e^{-\Omega_d(n)} \quad \text{for every } d\text{-local } D_c.$$

$D_q$ is essentially: sample $x \in \{0,1\}^n$ with bias $1/4$, then

if $|x|$ is odd, take $y$ uniform; otherwise take $y$ uniform conditioned on $|y| \equiv |x|/2 \bmod 2$.

# Plan

## Introduction and applications

Data structure lower bounds

Input-independent quantum advantage

## Proof outline

Granularity, or Hamming weight concentration

The graph pruning theorem

Classifying locally sampleable symmetric distributions

## Summary and open problems

# Proof overview

01 How to rule out uniform over weight exactly $n/3$

granularity

02 How to rule out uniform over weight $\geq n/2$

no sharp cutoff

03 The general case

**Overall idea.**
Assume $f(U)$ is close to the target, find many **independent** pieces of the output, and accumulate error across them.

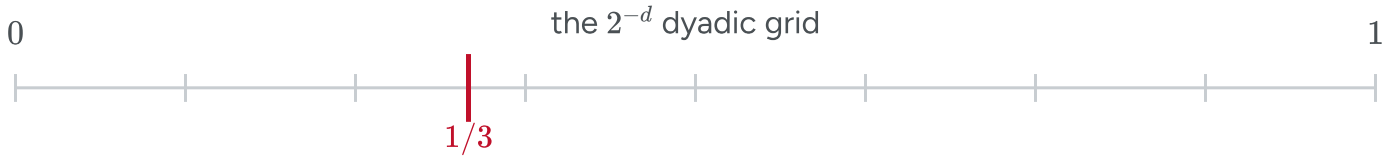
# Weight $n/3$

$f : \{0,1\}^m \to \{0,1\}^n$ is $d$-local
$f(U)$ is its output distribution on uniform input

Why can't $f(U)$ be uniform over $n$-bit strings of weight $n/3$?

marginals should be $1/3$-biased

marginals *are* multiples of $2^{-d}$




So every output bit is $\Omega(2^{-d})$ away from where it needs to be.

# Accumulating the error

$f : \{0,1\}^m \to \{0,1\}^n$ is $d$-local
$f(U)$ is its output distribution on uniform input

By granularity, each output bit contributes $\Omega_d(1)$ distance

By concentration, $r$ *independent* output bits contribute $1 - 2^{-\Omega_d(r)}$ distance

**Graph pruning theorem.**
By conditioning on $o_d(r)$ input bits, we obtain $r = \Omega_d(n)$ many independent output bits.

Let $S$ be the set of input bits given by the pruning theorem.

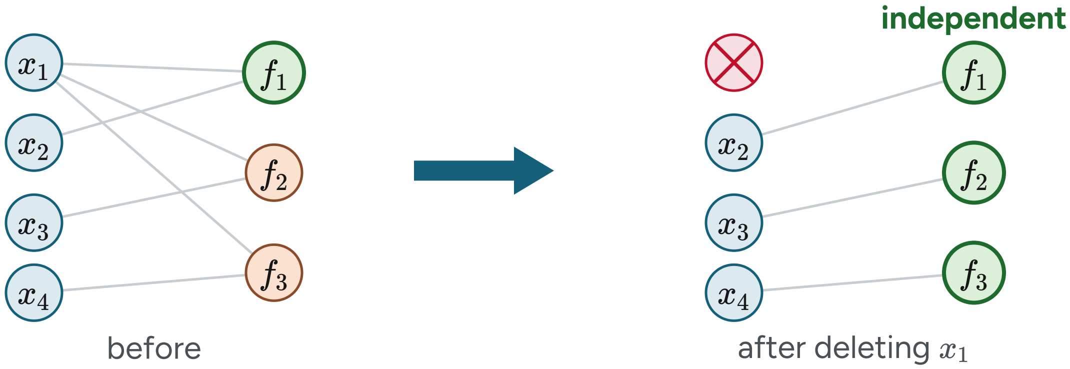
After conditioning on $x_S$ we get $1 - e^{-\Omega_d(n)}$ distance, and we can afford a union bound over all $2^{|S|}$ restrictions

$$\Delta_{\mathrm{TV}}\big(f(U),\, D_{n/3}\big) \;\geq\; 1 - 2^{|S| - \Omega_d(r)} \;=\; 1 - e^{-\Omega_d(n)}$$

# The pruning theorem

$f : \{0,1\}^m \to \{0,1\}^n$ is $d$-local
$f(U)$ is its output distribution on uniform input

Dependency graph $G = ([m], [n], E)$
Outputs $i, i'$ are **independent** when $N(i) \cap N(i') = \emptyset$




**Theorem.** If $G$ has right-degree $d$ and $n \gg_d 1$, there are $r \geq n^{0.9}$ and left vertices $S \subseteq [m]$ with $|S| \leq r/\log\log n$ such that $G \setminus S$ has $r$ independent right vertices.

# Proof of the pruning theorem

**Theorem.** If $G$ has right-degree $d$ and $n \gg_d 1$, there are $r \geq n^{0.9}$ and left vertices $S \subseteq [m]$ with $|S| \leq r/\log\log n$ such that $G \setminus S$ has $r$ independent right vertices.

For each $k \leq n^{0.01}$, let $S_k$ be the left vertices of degree $\geq k$

Greedily, $G \setminus S_k$ has $r_k \geq n/(dk)$ independent right vertices $r_k \geq n^{0.9}$

If the theorem fails, then $n/(dk) \leq r_k \leq |S_k| \cdot \log\log n$ for every such $k$

$$d \cdot n \;=\; |E| \;\geq\; \sum_{k=1}^{n^{0.01}} |S_k| \;\geq\; \sum_{k=1}^{n^{0.01}} \frac{n}{d\log\log n} \cdot \frac{1}{k} \;\geq\; \frac{n}{d\log\log n} \cdot 0.01\log n \;>\; d \cdot n$$

# Put things together

deleting left vertices $S$ $\iff$ conditioning on the input bits $x_S$

For $\rho \in \{0,1\}^S$, write $f|_\rho$ for the map with $x_S$ fixed to $\rho$. Then $f(U) \;=\; 2^{-|S|} \sum_{\rho \in \{0,1\}^S} f|_\rho(U)$.

After deleting $|S| \ll_d r$ left vertices, $G \setminus S$ has $r = \Omega_d(n)$ independent right vertices

$$
\begin{aligned}
\Delta_{\mathrm{TV}}\big(f(U),\ \mathrm{Ber}(1/3)^n\big) \;&\geq\; 1 - \sum_\rho \Big(1 - \Delta_{\mathrm{TV}}\big(f|_\rho(U),\ \mathrm{Ber}(1/3)^n\big)\Big) \\
&\geq\; 1 - \sum_\rho \Big(1 - \Delta_{\mathrm{TV}}\big(\textstyle\prod_{i \leq r} f_i|_\rho(U),\ \mathrm{Ber}(1/3)^r\big)\Big) \\
&\geq\; 1 - \sum_\rho 2^{-\Omega_d(r)} \;=\; 1 - 2^{|S| - \Omega_d(r)} \\
&\geq\; 1 - 2^{-\Omega_d(n)}
\end{aligned}
$$

# Weight $\geq n/2$

$f : \{0,1\}^m \to \{0,1\}^n$ is $d$-local
$f(U)$ is its output distribution on uniform input

Why can't $f(U)$ be uniform over $n$-bit strings of weight $\geq n/2$?

marginals should be $1/2$-biased

— and $1/2$ is on the dyadic grid.

**Sharp cutoff.**
$f(U)$ cannot build enough correlation to separate weight $< n/2$ from $\geq n/2$.

# Proof sketch

$f : \{0,1\}^m \to \{0,1\}^n$ is $d$-local
$f(U)$ is its output distribution on uniform input

**Graph pruning theorem.** Prune few input bits to get $r = \Omega_d(n)$ independent output bits.

Split the output bits by their marginal $\gamma_i = \Pr[f_i(U) = 1]$

## Granularity error

If most $\gamma_i$ are far from $1/2$, then $|\gamma_i - 1/2| \geq \Omega_d(1)$, and they accumulate $1 - e^{-\Omega_d(r)}$ error

## Concentration error

If most $\gamma_i$ are close to $1/2$, then $|f(U)|$ fluctuates by $\Theta_d(\sqrt{r})$ and cannot stay on one side of the cutoff point $n/2$

Independent output bits can be correlated with other bits
May not exhibit Gaussian-like concentration

**Large distance either way**

# A counterexample

Let $Z_1, \ldots, Z_{n/2}$ be independent unbiased bits and output

$Z_1,\ \neg Z_1$ | $Z_2,\ \neg Z_2$ | $Z_3,\ \neg Z_3$ | ... | $Z_{n/2},\ \neg Z_{n/2}$

Many independent output bits — but the total weight is always exactly $n/2$.

**The problem.**
Fluctuation of single bit can be cancelled by another output sharing the same input.

**The fix.**
Each ***neighborhood*** $(Z_i, \neg Z_i)$ is far from the correct marginal, so it has marginal error instead.

# Upgraded proof sketch

$f : \{0,1\}^m \to \{0,1\}^n$ is $d$-local
$f(U)$ is its output distribution on uniform input

**Upgraded graph pruning theorem.**
Prune few input bits to get $r = \Omega_d(n)$ independent output **neighborhoods**.

Split the output neighborhoods by their marginals

## Granularity error

If most are far from unbiased, then they accumulate large marginal error

## Concentration error

If most are close to unbiased, then $|f(U)|$ fluctuates by $\Theta_d(\sqrt{r})$ and cannot stay on one side of the cutoff point $n/2$

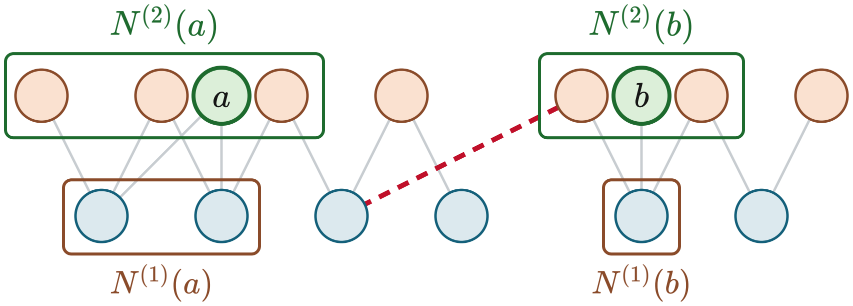
## Large distance either way

# Two-step independence

$N^{(1)}(i)$ = inputs that affect $f_i$

$N^{(2)}(i)$ = outputs that share an input with $f_i$

$N^{(2)}(i)$ is every output that can react when $f_i$ changes




$a$ and $b$ are **two-step independent** when $N^{(1)}(a') \cap N^{(1)}(b') = \emptyset$ for all $a' \in N^{(2)}(a)$, $b' \in N^{(2)}(b)$

**Upgraded pruning theorem.**
By removing few input bits, we obtain many two-step independent outputs with small $N^{(2)}$

Disjoint total input supports, so their whole blocks are independent

# Upgraded pruning theorem

**Theorem.** If $G$ has right-degree $d$ and $n \gg_d 1$, there are $r = \Omega_d(n)$ and left vertices $S \subseteq [m]$ with $|S| \ll_d r$ such that $G \setminus S$ has $r$ two-step independent right vertices with small $N^{(2)}$.

**Suppose the theorem fails**

- $\Rightarrow$ **we can delete $o_d(n)$ left vertices to remove $\Omega_d(n)$ edges**
- $\Rightarrow$ iterating, we delete $m = o_d(n)$ left vertices and reach the empty graph
- $\Rightarrow$ but then $S = [m]$ leaves all of $[n]$ two-step independent
- $\Rightarrow$ the theorem holds after all.

# Iterative edge removal

Call a right vertex $i$ **good** if every $j \in N^{(1)}(i)$ has degree $\leq k = \omega(1)$

**good** $\Rightarrow$ small $N^{(2)}$

If pruning $o(n)$ left vertices never yields $\Omega(n)$ two-step independent **good** right vertices, then we can delete $o(n)$ left vertices to remove $\Omega(n)$ edges.

*otherwise the upgraded pruning theorem holds*

## Case 1 · at least $n/2$ right vertices are bad

Let $T$ be the left vertices of degree $\geq k$. Every **bad** vertex touches at least one, so

$$n/2 \;\leq\; \#\mathrm{bad} \;\leq\; \sum_{j \in T} \deg(j)$$

and $\deg(j) \geq k$ for every $j \in T$, so $|T| \leq n/k$

Pruning $\leq n/k = o(n)$ left vertices removes $n/2$ edges. ✓

# Iterative edge removal

Call a right vertex $i$ **good** if every $j \in N^{(1)}(i)$ has degree $\leq k = \omega(1)$

**good** $\Rightarrow$ small $N^{(2)}$

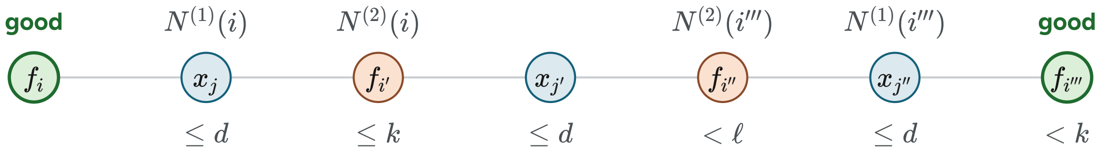
If pruning $o(n)$ left vertices never yields $\Omega(n)$ two-step independent **good** right vertices,
then we can delete $o(n)$ left vertices to remove $\Omega(n)$ edges.

*otherwise the upgraded pruning theorem holds*

## Case 2 · at least $n/2$ right vertices are good

Let $S_\ell$ be the left vertices of degree $\geq \ell$

Removing $S_\ell$ gives $r_\ell = (n/2)/(d^3k^2\ell)$ two-step independent **good** right vertices




By the failure assumption, $|S_\ell| \gg r_\ell \gg n/(k^2\ell)$

$$d \cdot n \;\geq\; |E| \;\geq\; \sum_{\ell \ll 2^{dk^2}} |S_\ell| \;\gg\; \sum_{\ell \ll 2^{dk^2}} \frac{n}{k^2} \cdot \frac{1}{\ell} \;\gg\; n \cdot \frac{dk^2}{k^2} \;>\; d \cdot n$$

# Weight $\geq n/2$: the full argument

Conditioning on $o(n)$ input bits gives $\Omega(n)$ independent output **neighborhoods**.

most are **far** from the marginal

each contributes $\Omega(1)$ distance

$1 - e^{-\Omega(n)}$ distance by concentration

most are **close** to the marginal

the weight distribution is nearly Gaussian

a Gaussian cannot produce a sharp cutoff

**Large distance either way**

# Uniform symmetric distributions

For $\emptyset \neq \Psi \subseteq \{0, \ldots, n\}$, write $D_\Psi = \mathrm{Unif}\{\, x \in \{0,1\}^n : |x| \in \Psi \,\}$
Suppose $f(U)$ is close to $D_\Psi$

## Tail regime

$|w - n/2| \geq \delta n$ for **every** $w \in \Psi$

the support avoids the middle entirely,
which has marginal / cutoff issues

## Central regime

$|w - n/2| \leq \delta n$ for **some** $w \in \Psi$

the support reaches the middle,
which enforces unbiased marginals

| $f(U) \equiv 0^n$ | $f(U) \equiv 1^n$ | $f(U) \sim \{0^n, 1^n\}$ | $\lvert f(U)\rvert \equiv 0 \bmod 2$ | $\lvert f(U)\rvert \equiv 1 \bmod 2$ | $f(U) \sim \{0,1\}^n$ |
|---|---|---|---|---|---|
| `zeros` | `ones` | `zeros-or-ones` | `evens` | `odds` | `evens-or-odds` |

# Why mod 2 is special

$\mathrm{Bin}(r, p) \bmod c$ has the same distribution as $\mathrm{Bin}(r, p) + 1 \bmod c$

only when $p = 1/2$ and $c = 2$

$r$ is the size of a two-step neighborhood $N^{(2)}(i)$

$p$ is the correct marginal of the neighborhood $N^{(2)}(i)$

$+1$ is for the independence of flipping output bit $i$

$\bmod c$ is to check continuity of the Hamming weight distribution modulo different numbers

# Dropping uniformity

**Theorem**

If a $d$-local $f(U)$ is close to **any** symmetric distribution, it is close to a mixture of

$$\texttt{evens},\ \texttt{odds},\ \text{and}\ \mathrm{Ber}(\gamma)^n \ \text{ with } \gamma \text{ an integer multiple of } 2^{-d}$$

and the mixing weights are biases of $\mathbf{F}_2$-polynomials of $O_d(1)$ degree with $O_d(n)$ monomials.

Condition on the high-degree input bits: $f(U)$ becomes a dyadic mixture of $P_1, P_2, \ldots$

each $P_i$ has small weight-variance

⇒ $P_i$ must *locally* look like $\mathrm{Ber}(\gamma_i)^n$ for a dyadic $\gamma_i$

$\gamma_i \neq 1/2$

$P_i$ is genuinely close to $\mathrm{Ber}(\gamma_i)^n$

$\gamma_i = 1/2$

$P_i$ is close to a structured mixture of **evens** and **odds**

# Plan

## Introduction and applications

Data structure lower bounds

Input-independent quantum advantage

## Proof outline

Granularity, or Hamming weight concentration

The graph pruning theorem

Classifying locally sampleable symmetric distributions

## Summary and open problems

# Summary

**Locally sampleable** (uniform) **symmetric distributions**

## Applications

Data structure lower bounds

Provable quantum advantages

## Techniques

Graph pruning theorems

Marginal / concentration error

### Open question I

What local distributions are *exact* symmetric?

### Open question II

Our analysis has a $\mathrm{tower}(d)$ loss.
What is the right quantitative bound?

# Open problem I: exact symmetric distributions

Our analysis has an exponentially small error term that persists even when the sampling is *exact*.

If a local distribution is *exactly* symmetric, what can it be?

**A concrete example**

$\mathrm{Ber}(1/4)^n + 2^{-n-1} \cdot$ **evens** $- 2^{-n-1} \cdot$ **odds** $\Leftrightarrow$ bitwise **AND** of **evens** and a uniform string

**Conjecture**

It is a mixture of bitwise post-processings of constantly many copies of **evens** and **odds**, with mixing weights given by low-degree $\mathbf{F}_2$-polynomials with linearly many monomials.

# Open problem II: how large must $n$ be?

Our classifications require the number of output bits $n$ to be **tower-ly larger** than the locality $d$.

Does the classification already hold for $n \geq 2^{\mathrm{poly}(d)}$?

### Where the tower comes from

Remove a constant fraction of edges costs
a telescoping sum below an exponential threshold

→

$\Omega(d)$ iterations of exponentiation produce
a $\mathrm{tower}(d)$ requirement

**Remark**

The tower-type dependence is necessary for the upgraded graph pruning itself.

For $n \ll \mathrm{tower}(d)$, there exists $G$ with right-degree $d$ such that removing any left-vertex set $S$ produces at most $|S|$ two-step independent right vertices.

# Thank you